

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/680,514

TIME: 10:17:17

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Yokoi, Haruhiko

7 Shiotsu, Yukimasa

8 Konishi, Noboru

10 (ii) TITLE OF INVENTION: NOVEL POLYPEPTIDES

12 (iii) NUMBER OF SEQUENCES: 45

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: NIXON & VANDERHYE P.C.

16 (B) STREET: 1100 North Glebe Rd. 8th floor

17 (C) CITY: Arlington

18 (D) STATE: VA

19 (E) COUNTRY: USA

20 (F) ZIP: 22201-4741

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/680,514

C--> 30 (B) FILING DATE: 06-Oct-2000

31 (C) CLASSIFICATION:

37 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/765,337

35 (B) FILING DATE:

38 (A) APPLICATION NUMBER: JP P.HEI.7-102625

39 (B) FILING DATE: 26-APR-1995

41 (viii) ATTORNEY/AGENT INFORMATION:

42 (A) NAME: Crawford, Arthur R.

43 (B) REGISTRATION NUMBER: 25,327

44 (C) REFERENCE/DOCKET NUMBER: 249-89

46 (ix) TELECOMMUNICATION INFORMATION:

47 (A) TELEPHONE: 703-816-4000

48 (B) TELEFAX: 703-816-4100

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 328 amino acids

55 (B) TYPE: amino acid

56 (C) STRANDEDNESS:

57 (D) TOPOLOGY: linear

59 (ii) MOLECULE TYPE: peptide

64 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

66 Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu

67 1 5 10 15

69 Arg Asp Ser His Val Leu His Ser Arg Leu Ser Gln Cys Pro Glu Val

70 20 25 30

ENTERED

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/680,514

TIME: 10:17:17

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

```

72 His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala Val Asp Phe Ser Leu
73      35      40      45
75 Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala Gln Asp Ile Leu
76      50      55      60
78 Gly Ala Val Thr Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln
79      65      70      75      80
81 Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln
82      85      90      95
84 Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr Gln Leu
85      100      105      110
87 Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile Phe
88      115      120      125
90 Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met Leu
91      130      135      140
93 Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala Pro Thr Tyr Arg Ala
94      145      150      155      160
96 Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser Leu Glu Gln Val Arg
97      165      170      175
99 Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu Lys Leu Cys Ala Thr
100      180      185      190
102 Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu Leu Gly His Ser Leu
103      195      200      205
105 Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro Ser Gln Ala Leu Gln
106      210      215      220
108 Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly Leu Phe Leu Tyr Gln
109      225      230      235      240
111 Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro Glu Leu Gly Pro Thr
112      245      250      255
114 Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe Ala Thr Thr Ile Trp
115      260      265      270
117 Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala Leu Gln Pro Thr Gln
118      275      280      285
120 Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln Arg Arg Ala Gly Gly
121      290      295      300
123 Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr Arg
124      305      310      315      320
126 Val Leu Arg His Leu Ala Gln Pro
127      325

```

129 (2) INFORMATION FOR SEQ ID NO: 2:

131 (i) SEQUENCE CHARACTERISTICS:

132 (A) LENGTH: 340 amino acids

133 (B) TYPE: amino acid

134 (C) STRANDEDNESS:

135 (D) TOPOLOGY: linear

137 (ii) MOLECULE TYPE: peptide

142 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

144 Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu
145 1      5      10      15
147 Arg Asp Ser His Val Leu His Ser Arg Leu Ser Gln Cys Pro Glu Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/680,514

DATE: 12/17/2001

TIME: 10:17:17

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

```

148          20          25          30
150 His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala Val Asp Phe Ser Leu
151          35          40          45
153 Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala Gln Asp Ile Leu
154          50          55          60
156 Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln
157          65          70          75          80
159 Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln
160          85          90          95
162 Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr Gln Leu
163          100          105          110
165 Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile Phe
166          115          120          125
168 Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met Leu
169          130          135          140
171 Val Gly Gly Ser Thr Leu Cys Val Arg Gly Gly Gly Ser Gly Gly Gly
172          145          150          155          160
174 Ser Gly Gly Gly Ser Arg Ala Pro Thr Tyr Arg Ala Ser Ser Leu Pro
175          165          170          175
177 Gln Ser Phe Leu Leu Lys Ser Leu Glu Gln Val Arg Lys Ile Gln Gly
178          180          185          190
180 Asp Gly Ala Ala Leu Gln Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys
181          195          200          205
183 His Pro Glu Glu Leu Val Leu Leu Gly His Ser Leu Gly Ile Pro Trp
184          210          215          220
186 Ala Pro Leu Ser Ser Cys Pro Ser Gln Ala Leu Gln Leu Ala Gly Cys
187          225          230          235          240
189 Leu Ser Gln Leu His Ser Gly Leu Phe Leu Tyr Gln Gly Leu Leu Gln
190          245          250          255
192 Ala Leu Glu Gly Ile Ser Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu
193          260          265          270
195 Gln Leu Asp Val Ala Asp Phe Ala Thr Thr Ile Trp Gln Gln Met Glu
196          275          280          285
198 Glu Leu Gly Met Ala Pro Ala Leu Gln Pro Thr Gln Gly Ala Met Pro
199          290          295          300
201 Ala Phe Ala Ser Ala Phe Gln Arg Arg Ala Gly Gly Val Leu Val Ala
202          305          310          315          320
204 Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr Arg Val Leu Arg His
205          325          330          335
207 Leu Ala Gln Pro
208          340

```

210 (2) INFORMATION FOR SEQ ID NO: 3:

212 (i) SEQUENCE CHARACTERISTICS:

213 (A) LENGTH: 344 amino acids

214 (B) TYPE: amino acid

215 (C) STRANDEDNESS:

216 (D) TOPOLOGY: linear

218 (ii) MOLECULE TYPE: peptide

223 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/680,514

TIME: 10:17:17

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

```

225 Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu
226 1 5 10 15
228 Arg Asp Ser His Val Leu His Ser Arg Leu Ser Gln Cys Pro Glu Val
229 20 25 30
231 His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala Val Asp Phe Ser Leu
232 35 40 45
234 Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala Gln Asp Ile Leu
235 50 55 60
237 Gly Ala Val Thr Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln
238 65 70 75 80
240 Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln
241 85 90 95
243 Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr Gln Leu
244 100 105 110
246 Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile Phe
247 115 120 125
249 Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met Leu
250 130 135 140
252 Val Gly Gly Ser Thr Leu Cys Val Arg Ser Gly Gly Gly Ser Gly Gly
253 145 150 155 160
255 Gly Ser Gly Gly Gly Ser Gly Gly Gly Arg Ala Pro Thr Tyr Arg Ala
256 165 170 175
258 Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser Leu Glu Gln Val Arg
259 180 185 190
261 Lys Ile Gln Gly Asp Gly Ala Ala Leu Gln Glu Lys Leu Cys Ala Thr
262 195 200 205
264 Tyr Lys Leu Cys His Pro Glu Leu Val Leu Leu Gly His Ser Leu
265 210 215 220
267 Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro Ser Gln Ala Leu Gln
268 225 230 235 240
270 Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly Leu Phe Leu Tyr Gln
271 245 250 255
273 Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro Glu Leu Gly Pro Thr
274 260 265 270
276 Leu Asp Thr Leu Gln Leu Asp Val Ala Asp Phe Ala Thr Thr Ile Trp
277 275 280 285
279 Gln Gln Met Glu Glu Leu Gly Met Ala Pro Ala Leu Gln Pro Thr Gln
280 290 295 300
282 Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln Arg Arg Ala Gly Gly
283 305 310 315 320
285 Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr Arg
286 325 330 335
288 Val Leu Arg His Leu Ala Gln Pro
289 340

```

291 (2) INFORMATION FOR SEQ ID NO: 4:

293 (i) SEQUENCE CHARACTERISTICS:

294 (A) LENGTH: 1047 base pairs

295 (B) TYPE: nucleic acid

296 (C) STRANDEDNESS: double

RAW SEQUENCE LISTING

DATE: 12/17/2001

PATENT APPLICATION: US/09/680,514

TIME: 10:17:17

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

```

297      (D) TOPOLOGY: linear
299      (ii) MOLECULE TYPE: other nucleic acid
300      (A) DESCRIPTION: /desc = "synthetic DNA"
303      (ix) FEATURE:
304          (A) NAME/KEY: sig_peptide
305          (B) LOCATION: 1..63
307      (ix) FEATURE:
308          (A) NAME/KEY: mat_peptide
309          (B) LOCATION: 64..1047
311      (ix) FEATURE:
312          (A) NAME/KEY: CDS
313          (B) LOCATION: 1..1047
316      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
318 ATG GAG CTG ACT GAA TTG CTC CTC GTG GTC ATG CTT CTC CTA ACT GCA      48
319 Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala
320 -21 -20                      -15                      -10
322 AGG CTA ACG CTG TCC AGC CCG GCT CCT CCT GCT TGT GAC CTC CGA GTC      96
323 Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val
324 -5                      1                      5                      10
326 CTC AGT AAA CTG CTT CGT GAC TCC CAT GTC CTT CAC AGC AGA CTG AGC      144
327 Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser
328          15                      20                      25
330 CAG TGC CCA GAG GTT CAC CCT TTG CCT ACA CCT GTC CTG CTG CCT GCT      192
331 Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala
332          30                      35                      40
334 GTG GAC TTT AGC TTG GGA GAA TGG AAA ACC CAG ATG GAG GAG ACC AAG      240
335 Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys
336          45                      50                      55
338 GCA CAG GAC ATT CTG GGA GCA GTG ACC CTT CTG CTG GAG GGA GTG ATG      288
339 Ala Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met
340 60                      65                      70                      75
342 GCA GCA CGG GGA CAA CTG GGA CCC ACT TGC CTC TCA TCC CTC CTG GGG      336
343 Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu Gly
344          80                      85                      90
346 CAG CTT TCT GGA CAG GTC CGT CTC CTC CTT GGG GCC CTG CAG AGC CTC      384
347 Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln Ser Leu
348          95                      100                      105
350 CTT GGA ACC CAG CTT CCT CCA CAG GGC AGG ACC ACA GCT CAC AAG GAT      432
351 Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys Asp
352          110                      115                      120
354 CCC AAT GCC ATC TTC CTG AGC TTC CAA CAC CTG CTC CGA GGA AAG GTG      480
355 Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val
356          125                      130                      135
358 CGT TTC CTG ATG CTT GTA GGA GGG TCC ACC CTC TGC GTA CGG CGG GCG      528
359 Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala
360 140                      145                      150                      155
362 CCA ACA TAT CGC GCC TCG AGT CTA CCA CAG AGC TTC CTT TTA AAA AGC      576
363 Pro Thr Tyr Arg Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Ser
364          160                      165                      170

```

VERIFICATION SUMMARY

DATE: 12/17/2001

PATENT APPLICATION: US/09/680,514

TIME: 10:17:18

Input Set : N:\Crf3\RULE60\09680514.txt

Output Set: N:\CRF3\12172001\I680514.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:1502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39